

Abstract

In a randomly oriented graph containing vertices x and y , denote by $\{x \rightarrow y\}$ the event that there is a directed path from x to y . We study the correlation between the events $\{x \rightarrow y\}$ and $\{y \rightarrow z\}$ for a (large) oriented complete bipartite graph with orientation chosen uniformly at random. We classify the cases of positive and negative correlation respectively in terms of the relative proportions of the sizes of the color classes of the graph.

Path correlations in a randomly oriented complete bipartite graph

Erik Aas

1 Introduction

Let G be an arbitrary finite graph whose edges e all have been assigned probabilities $p(e)$. We get a random graph by including the edge e in the graph with probability $p(e)$ independently of everything else. Let $\{x \leftrightarrow y\}$ denote the event that there is a path connecting the two nodes x and y in this random graph.

In [4] it was observed that when all $p(e) = \frac{1}{2}$, the probability $P(x \leftrightarrow y)$ coincided with the probability of $\{x \rightarrow y\}$, the event of there being a directed path from x to y in a uniformly chosen random orientation of the edges of G (so here we are comparing the probabilities of two different events in two different probability spaces). Clearly choosing an orientation uniformly at random amounts to orient each edge either way with equal probability $1/2$, independently of all other edges.

A natural question, then, is what other properties those two probability spaces share.

It is a basic fact [3] that in the undirected random graph model defined above, for any increasing events A and B (like $\{x \leftrightarrow z\}$ and $\{z \leftrightarrow y\}$),

$$P(A \cap B) \geq P(A)P(B),$$

that is, A and B are positively (or, to be pedantic, nonnegatively) correlated.

An analogue of this fact, stated in [4], is that in any *randomly directed* graph, the events $\{x \rightarrow y\}$ and $\{y \rightarrow z\}$ are positively correlated, that is,

$$P(x \rightarrow y \rightarrow z) \geq P(x \rightarrow y)P(y \rightarrow z).$$

This motivates the following question: *Let G be a graph containing three vertices x, y, z . Does*

$$P(x \rightarrow y \rightarrow z) \geq P(x \rightarrow y)P(y \rightarrow z)$$

hold, that is, are $\{x \rightarrow y\}$ and $\{y \rightarrow z\}$ positively correlated? Here, and in the following, we write $\{x \rightarrow y \rightarrow z\} := \{x \rightarrow y\} \cap \{y \rightarrow z\}$.

Obviously this depends on the graph G , as it is easy to find graphs for which $P(x \rightarrow y \rightarrow z) - P(x \rightarrow y)P(y \rightarrow z)$ has any given sign (including 0), and a simple characterisation of all graphs with, say, positive correlation seems hard to find. It is known [1] that for the complete graph K_n , this quantity is negative for $n = 3$, zero for $n = 4$ and positive for

all $n > 4$. For a slightly different model, in [2] it is shown that when the graph is 'dense', the analogous correlation is positive. Here we will study the same correlation between the events $\{x \rightarrow y\}$ and $\{y \rightarrow z\}$ in a uniformly chosen orientation of the edges of the complete bipartite graph $K_{m,n}$.

2 Result

Throughout the remainder of this note, let $A = \{x \rightarrow y\}$ and $B = \{y \rightarrow z\}$. We denote the complement of a set (or an event) A by A^c . We define $\{x \not\rightarrow y\}$ to be $\{x \rightarrow y\}^c$, that is, there is not path from x to y . As a technical convenience the object of study will be $RC_{m,n} := \frac{P(A^c \cap B^c) - P(A^c)P(B^c)}{P(A^c \cap B^c)}$, the relative covariance between A^c and B^c , rather than $P(A \cap B) - P(A)P(B)$. Observe that $P(A \cap B) - P(A)P(B) = P(A^c \cap B^c) - P(A^c)P(B^c)$ - this holds for any two events A, B . In particular, A and B are positively correlated if and only if $RC_{m,n}$ is positive. Observe that the relative covariance can be rewritten as (and this might be the more convenient way of thinking about it) $RC_{m,n} = 1 - \frac{P(A^c)P(B^c)}{P(A^c \cap B^c)} = 1 - \frac{P(B^c)}{P(B^c|A^c)}$.

We fix some more notation:

- The nodes of $K_{m,n}$ are partitioned into two sets X and Y of sizes m and n respectively.
- $m = \lfloor \beta n \rfloor$ for some fixed positive constant β .
- $m, n \geq 2$.
- a, b, c, d are four distinct vertices of $K_{m,n}$; the first three belong to X and d belongs to Y .
- The limit $\lim_{n \rightarrow \infty} RC_{m,n}$ is denoted by RC .

Theorem 2.1. *The value of RC is given by the following table.*

X	Y	$\beta < 1$	$\beta = 1$	$\beta > 1$
x, y, z		-1/3	-1/3	-1/3
x, y	z	1/2	1/5	-1
x, z	y	1	1/5	0

Table 1: The relative covariance between $\{x \not\rightarrow y\}$ and $\{y \not\rightarrow z\}$, according to which partition the vertices belong to, and to the proportion β of the number X -vertices to the number of Y -vertices.

We see that letting $m = \lfloor \beta n \rfloor$ for a fixed constant β is not as restrictive as might seem at first thought.

3 Proof

The proof of Theorem 2.1 will follow from a number of lemmas estimating the probabilities $P(A^c), P(B^c), P(A^c \cap B^c)$ in terms of n . A common

feature of these estimates is that the lower bounds, which are trivial to obtain, are close to the harder-to-prove upper bounds. If S and T are two disjoint sets of vertices in $K_{m,n}$, an ' ST -witness' is defined to be a vertex u for which there is at least one edge from S to u and at least one edge from u to T .

For a vertex s in $K_{m,n}$, the set O_s will loosely be defined as the set of vertices in $X' \cup Y'$ (thus in X' if $a \in Y$ and in Y' if $a \in X$) which can be reached in exactly one step from s , X' and Y' being defined separately in each section where this notation is used. We denote $|X'|$ and $|Y'|$ by m' and n' . The set I_a of vertices which reach a in exactly one step is similarly defined.

Estimating these probabilities will be a lot of repetitive work. The following inequality will be used several times: if $s, t \geq \alpha$, then $st \geq \alpha s + \alpha t - \alpha^2$.

Below, when summing over subsets of nodes denoted by upper case letters, the sizes of these sets will often be denoted by the corresponding lower case letters.

To estimate sums of the form $\sum_{s=0}^n \sum_{t=0}^{n-s} \binom{n}{s} \binom{n-s}{t} \left(\frac{1}{2}\right)^{st}$ we will split them into several parts according to whether $s \geq \alpha$ or $t \geq \alpha$ for some suitably chosen constant α (depending only on β).

(i) $P(b \nrightarrow a)$

Lemma 3.1.

$$P(b \nrightarrow a) \sim 2 \left(\frac{1}{2}\right)^n.$$

Proof. Let $X' = X - \{a, b\}$, $Y' = Y$.

A lower bound is given by $P(b \nrightarrow a) \geq P(\{\text{there is no edge directed away from } b\} \cup \{\text{there is no edge directed towards } a\}) = 2 \left(\frac{1}{2}\right)^n - \left(\frac{1}{2}\right)^{2n}$, by inclusion-exclusion.

By calculating the probability that there is no path from b to a of length at most 4, we get the following upper bound: $P(b \nrightarrow a) = \sum_{S, T \subseteq Y} P(b \nrightarrow a | O_b = S, I_a = T) P(O_b = S, I_a = T) \leq \left(\frac{1}{2}\right)^{2n} \sum_{S, T \subseteq Y: S \cap T = \emptyset} P(\text{no } x \in X' \text{ is an } ST\text{-witness}) = \left(\frac{1}{2}\right)^{2n} \sum_{s=0}^n \sum_{t=0}^{n-s} \binom{n}{s} \binom{n-s}{t} \left(\left(\frac{1}{2}\right)^s + \left(\frac{1}{2}\right)^t - \left(\frac{1}{2}\right)^{s+t}\right)^{m-2}.$

Note that the partial sum corresponding to $st = 0$ is equal to the lower bound. We now show that the other terms sum to $o\left(\left(\frac{1}{2}\right)^n\right)$. Split the remaining sum into the following four parts: S_1 : $s, t \geq \alpha$; S_2 : $1 \leq s \leq \alpha \leq t$; S_3 : $1 \leq t \leq \alpha \leq s$; S_4 : $1 \leq s, t \leq \alpha$.

Note that in the S_1 case, $\left(\frac{1}{2}\right)^s + \left(\frac{1}{2}\right)^t - \left(\frac{1}{2}\right)^{s+t} \leq \left(\frac{1}{2}\right)^{\alpha-1}$. Hence $S_1 \leq \left(\frac{1}{2}\right)^{2n} \sum_{s=\alpha}^n \binom{n}{s} \sum_{t=\alpha}^{n-s} \binom{n-s}{t} \left(\frac{1}{2}\right)^{(\alpha-1)(m-2)} = \left(\frac{1}{2}\right)^{2n+(\alpha-1)(m-2)} 3^n = o\left(\left(\frac{1}{2}\right)^n\right)$, the last equality holding when choosing α large enough.

$S_2 \leq \left(\frac{1}{2}\right)^{2n} \sum_{s=1}^{\alpha} \binom{n}{s} \sum_{t=\alpha}^{n-s} \binom{n-s}{t} \left(\left(\frac{1}{2}\right)^s + \left(\frac{1}{2}\right)^t\right)^{m-2} \leq \left(\frac{1}{2}\right)^{2n} \sum_{s=1}^{\alpha} n^{\alpha} \sum_{t=0}^n \binom{n}{t} \left(\left(\frac{1}{2}\right) + \left(\frac{1}{2}\right)^{\alpha}\right)^{m-2} \leq \left(\frac{1}{2}\right)^n \alpha n^{\alpha} \left(\left(\frac{1}{2}\right) + \left(\frac{1}{2}\right)^{\alpha}\right)^{m-2} = o\left(\left(\frac{1}{2}\right)^n\right)$, if $\alpha > 0$.

By symmetry, we may choose α possibly even larger so that $S_3 = o\left(\left(\frac{1}{2}\right)^n\right)$ holds.

Clearly, $S_4 = o\left(\left(\frac{1}{2}\right)^n\right)$.

Hence $P(b \nrightarrow a) - 2 \left(\frac{1}{2}\right)^n \leq S_1 + S_2 + S_3 + S_4 = o\left(\left(\frac{1}{2}\right)^n\right)$. \square

(ii) $P(d \nrightarrow a)$

Lemma 3.2. $P(d \nrightarrow a) \sim \left(\frac{1}{2}\right)^m + \left(\frac{1}{2}\right)^n$. (This is $\sim \left(\frac{1}{2}\right)^m$ for $\beta < 1$, $\sim 2\left(\frac{1}{2}\right)^n$ for $\beta = 1$, and $\sim \left(\frac{1}{2}\right)^n$ for $\beta > 1$.)

Proof. Let $X' = X - \{a\}$, $Y' = Y - \{d\}$. The probability is bounded from below by $P(d \nrightarrow a) \geq P(\text{no edge leaves } d \text{ or no edge enters } a) \geq \left(\frac{1}{2}\right)^m + \left(\frac{1}{2}\right)^n - \left(\frac{1}{2}\right)^{m+n-1}$.

For the upper bound, we calculate the probability that there is no path from d to a of length at most 3: $P(d \nrightarrow a) = \sum_{S \subseteq Y', T \subseteq X'} P(a \nrightarrow d | I_a = S, O_d = T) P(I_a = S, O_d = T) \leq \left(\frac{1}{2}\right)^{m+n-1} \sum_{s=0}^{n-1} \binom{n-1}{s} \sum_{t=0}^{m-1} \binom{m-1}{t} P(\text{no edge from } S \text{ to } T) = \left(\frac{1}{2}\right)^{m+n-1} \sum_{s=0}^{n-1} \binom{n-1}{s} \sum_{t=0}^{m-1} \binom{m-1}{t} \left(\frac{1}{2}\right)^{st}$. The partial sum with $st = 0$ equals the lower bound. We now show that the remaining terms sum to $o\left(\left(\frac{1}{2}\right)^m + \left(\frac{1}{2}\right)^n\right)$ by splitting their sum into the following cases: S_1 : $s, t \geq \alpha$, S_2 : $1 \leq s \leq \alpha$, and S_3 : $1 \leq t \leq \alpha$.

Using $s, t \geq \alpha \Rightarrow st \geq \alpha s + \alpha t - \alpha^2$, $S_1 = \left(\frac{1}{2}\right)^{m+n-1} \sum_{s=\alpha}^{n-1} \binom{n-1}{s} \sum_{t=\alpha}^{m-1} \binom{m-1}{t} \left(\frac{1}{2}\right)^{st} \leq \left(\frac{1}{2}\right)^{m+n-1} \sum_{s=\alpha}^{n-1} \binom{n-1}{s} \sum_{t=\alpha}^{m-1} \binom{m-1}{t} \left(\frac{1}{2}\right)^{\alpha s + \alpha t - \alpha^2} \leq \left(\frac{1}{2}\right)^{m+n-1} 2^{\alpha^2} \sum_{s=0}^{n-1} \binom{n-1}{s} \left(\frac{1}{2}\right)^{\alpha s} \sum_{t=0}^{m-1} \binom{m-1}{t} \left(\frac{1}{2}\right)^{\alpha t} \leq \left(\frac{1}{2}\right)^{m+n-1} 2^{\alpha^2} \left(1 + \left(\frac{1}{2}\right)^\alpha\right)^{m+n-2} = o\left(\left(\frac{1}{2}\right)^m + \left(\frac{1}{2}\right)^n\right)$, choosing α large enough. Similarly, $S_2 = \left(\frac{1}{2}\right)^{m+n-1} \sum_{s=1}^{\alpha} \binom{n-1}{s} \sum_{t=0}^{m-1} \binom{m-1}{t} \left(\frac{1}{2}\right)^{st} \leq \left(\frac{1}{2}\right)^{m+n-1} (n-1)^\alpha \alpha \sum_{t=0}^{m-1} \binom{m-1}{t} \left(\frac{1}{2}\right)^t = \alpha(n-1)^\alpha \left(\frac{1}{2}\right)^n \left(\frac{3}{4}\right)^{m-1} = o\left(\left(\frac{1}{2}\right)^m + \left(\frac{1}{2}\right)^n\right)$.

A similar argument shows $S_3 = o\left(\left(\frac{1}{2}\right)^m + \left(\frac{1}{2}\right)^n\right)$.

Hence $P(d \nrightarrow a) - \left(\left(\frac{1}{2}\right)^m + \left(\frac{1}{2}\right)^n\right) \leq S_1 + S_2 + S_3 = o\left(\left(\frac{1}{2}\right)^m + \left(\frac{1}{2}\right)^n\right)$. \square

(iii) $P(b \nrightarrow d \nrightarrow a)$

Lemma 3.3. $P(b \nrightarrow d \nrightarrow a) \sim 2\left(\frac{1}{2}\right)^{m+n-1} + \left(\frac{1}{2}\right)^{2n}$

Proof. Let $X' = X - \{a, b\}$, $Y' = Y - \{d\}$. For the lower bound, we calculate the probability $P(\{\text{the edge between } b \text{ and } d, \text{ and the edge between } d \text{ and } a, \text{ form a directed path from } a \text{ to } b\} \cap (\{O_b = O_d = \emptyset\} \cup \{O_b = I_a = \emptyset\} \cup \{I_d = I_a = \emptyset\}))$, which is $2\left(\frac{1}{2}\right)^{m+n-1} + \left(\frac{1}{2}\right)^{2n} - \left(\frac{1}{2}\right)^{m+2n-3}$, by inclusion-exclusion; $P(b \nrightarrow d \nrightarrow a) \geq 2\left(\frac{1}{2}\right)^{m+n-1} + \left(\frac{1}{2}\right)^{2n} - \left(\frac{1}{2}\right)^{m+2n-3}$.

To get a working upper bound, it is sufficient to calculate the probability of there being no path from b to d or from d to a , either of length at most 3. Conditioning on $I_a = S, O_b = T, I_d = U, O_d = V$, there may be no edge from T to U , nor from V to S . The edges $\{b, d\}$ and $\{a, d\}$ form a directed path from a to b . This implies that S and T must be disjoint.

Hence

$$P(b \nrightarrow d \nrightarrow a)$$

$$= \sum_{S, T, U, V} P(b \nrightarrow d \nrightarrow a | I_a = S, O_b = T, I_d = U, O_d = V) P(I_a = S, O_b = T, I_d = U, O_d = V)$$

$$\leq \left(\frac{1}{2}\right)^{m+2n-2} \sum_{s=0}^{n-1} \binom{n-1}{s} \sum_{t=0}^{n-1-s} \binom{n-1-s}{t} \sum_{u=0}^{m-2} \binom{m-2}{u} \left(\frac{1}{2}\right)^{tu+s(m-2-u)}.$$

The sum of the terms for which $s = t = 0$, $t = u = 0$ or $t = u - (m - 2) = 0$ equals the lower bound. The remaining sum is split into the following cases: S_1 : $s, t \geq \alpha$, S_2 : $1 \leq t \leq \alpha \leq s$, S_3 : $1 \leq s \leq \alpha \leq t$, and S_4 : $1 \leq s, t \leq \alpha$.

$$\begin{aligned} S_1 &\leq 2^{2\alpha^2} \left(\frac{1}{2}\right)^{2n+m-2} \sum_{s=\alpha}^{n-1} \binom{n-1}{s} \sum_{t=\alpha}^{n-1-s} \binom{n-1-s}{t} \sum_{u=0}^{m-2} \left(\frac{1}{2}\right)^{\alpha(m-2)} = \\ &\left(\frac{1}{2}\right)^{2n+m-2} 3^{n-1} \left(\frac{1}{2}\right)^{(\alpha-1)(m-2)} = o\left(\left(\frac{1}{2}\right)^{m+n-2} + \left(\frac{1}{2}\right)^{2n}\right). \\ S_2 &\leq \left(\frac{1}{2}\right)^{2n+m-2} \sum_{s=1}^{\alpha} \binom{n-1}{s} \sum_{t=\alpha}^{n-1-s} \binom{n-1-s}{t} \sum_{u=0}^{m-2} \left(\frac{1}{2}\right)^{tu+s(m-2-u)} \leq \\ &\left(\frac{1}{2}\right)^{2n+m-2} (n-1)^{\alpha} \alpha \sum_{t=\alpha}^{n-1} \binom{n-1}{t} \sum_{u=0}^{m-2} \binom{m-2}{u} \left(\frac{1}{2}\right)^{\alpha u+m-2-u} = \left(\frac{1}{2}\right)^{2(m-2+n)} (n-1)^{\alpha} \alpha 2^{n-1} \left(1 + \left(\frac{1}{2}\right)^{\alpha-1}\right)^{m-2} = o\left(\left(\frac{1}{2}\right)^{m+n-1} + \left(\frac{1}{2}\right)^{2n}\right). \end{aligned}$$

By symmetry with S_2 , we deduce $S_3 = o\left(\left(\frac{1}{2}\right)^{m+n-1} + \left(\frac{1}{2}\right)^{2n}\right)$.

$$S_4 \leq \left(\frac{1}{2}\right)^{2n+m-2} (n-1)^{2\alpha} \alpha^2 = o\left(\left(\frac{1}{2}\right)^{m+n-1} + \left(\frac{1}{2}\right)^{2n}\right).$$

We conclude that

$$P(b \nrightarrow d \nrightarrow a) \sim 2 \left(\frac{1}{2}\right)^{m+n-1} + \left(\frac{1}{2}\right)^{2n}.$$

□

(iv) $c \nrightarrow b \nrightarrow a$

Lemma 3.4.

$$P(c \nrightarrow b \nrightarrow a) \sim 3 \left(\frac{1}{2}\right)^{2n}$$

Proof. For the lower bound, note that $P(c \nrightarrow b \nrightarrow a) \geq P(\{O_c = O_b = \emptyset\} \cup \{O_c = I_a = \emptyset\} \cup \{I_b = I_a = \emptyset\}) \geq 3 \left(\frac{1}{2}\right)^{2n} - 2 \left(\frac{1}{2}\right)^{3n}$.

For the upper bound, we will sum over $U \subseteq Y'$, $V = Y' - U$, $S \subseteq U$, and $T \subseteq V$. When doing so, an expression for the probability that a given vertex $x' \in X'$ is not a TU -witness and not a VS -witness is needed. The probability of the complementary event is the probability of x' being a TU -witness or a VS -witness. The separate probabilities for these last two events are $P(x' \text{ is a } TU\text{-witness}) = \left(1 - \left(\frac{1}{2}\right)^{|T|}\right) \left(1 - \left(\frac{1}{2}\right)^{|U|}\right)$ and $P(x' \text{ is an } SV\text{-witness}) = \left(1 - \left(\frac{1}{2}\right)^{|S|}\right) \left(1 - \left(\frac{1}{2}\right)^{|V|}\right)$. The probability of their intersection is $\left(1 - \left(\frac{1}{2}\right)^{|S|}\right) \left(1 - \left(\frac{1}{2}\right)^{|T|}\right)$, using $S \subseteq U$, $T \subseteq V$. By inclusion-exclusion, we get $P(x \in X' \text{ is not a } TU\text{-witness, nor a } VS\text{-witness}) = 1 - \left(1 - \left(\frac{1}{2}\right)^{|T|}\right) \left(1 - \left(\frac{1}{2}\right)^{|U|}\right) - \left(1 - \left(\frac{1}{2}\right)^{|S|}\right) \left(1 - \left(\frac{1}{2}\right)^{|V|}\right) + \left(1 - \left(\frac{1}{2}\right)^{|S|}\right) \left(1 - \left(\frac{1}{2}\right)^{|T|}\right)$, which simplifies to $\left(\frac{1}{2}\right)^{|U|} + \left(\frac{1}{2}\right)^{|V|} - \left(\frac{1}{2}\right)^{|T|+|U|} - \left(\frac{1}{2}\right)^{|S|+|V|} + \left(\frac{1}{2}\right)^{|S|+|T|}$.

$$P(c \nrightarrow b \nrightarrow a)$$

$$= \sum_{S,T,U,V} P(c \nrightarrow b \nrightarrow a | I_a = S, O_c = T, I_b = U, O_b = V) P(I_a = S, O_c = T, I_b = U, O_b = V)$$

$$\leq \sum_{S,T,U,V} P(\text{no } x \in X' \text{ is a } TU\text{-witness, nor a } VS\text{-witness}) \left(\frac{1}{2}\right)^{3n}$$

$$= \left(\frac{1}{2}\right)^{3n} \sum_{u=0, u+v=n}^n \binom{n}{u} \sum_{s=0}^u \binom{u}{s} \sum_{t=0}^v \binom{v}{t} \left(\left(\frac{1}{2}\right)^u + \left(\frac{1}{2}\right)^v - \left(\frac{1}{2}\right)^{t+u} - \left(\frac{1}{2}\right)^{s+v} + \left(\frac{1}{2}\right)^{s+t} \right)^{m-3}.$$

The sum of the terms with $s = t = 0$ or $u = 0$ or $v = 0$ equals the lower bound. The other terms sum to $o((\frac{1}{2})^{2n})$, as we now turn to show. Since (u, t) and (v, s) are interchangeable, we need only consider the following cases: S_1 : $s, t \geq \alpha$; S_2 : $1 \leq s \leq \alpha \leq t, u$; S_3 : $1 \leq s \leq \alpha \leq t, 1 \leq u \leq \alpha$; S_4 : $1 \leq s, t \leq \alpha$.

$$\begin{aligned} S_1 &\leq \left(\frac{1}{2}\right)^{3n} \sum_{u=\alpha, u+v=n}^{n-\alpha} \binom{n}{u} \sum_{s=\alpha}^u \binom{u}{s} \sum_{t=0}^v \binom{v}{t} \left(\left(\frac{1}{2}\right)^u + \left(\frac{1}{2}\right)^v - \left(\frac{1}{2}\right)^{t+u} - \left(\frac{1}{2}\right)^{s+v} + \left(\frac{1}{2}\right)^{s+t} \right)^{m-3} \\ &\leq \left(\frac{1}{2}\right)^{3n} \sum_{u=0, u+v=n}^n \binom{n}{u} \sum_{s=0}^u \binom{u}{s} \sum_{t=0}^v \binom{v}{t} \left(\frac{1}{2}\right)^{(\alpha-2)(m-3)} = O\left(\left(\frac{1}{2}\right)^{(\beta(\alpha-2)+1)n}\right), \end{aligned}$$

which is $o((\frac{1}{2})^{2n})$ when choosing α large enough (e.g. $\alpha > 2(1 + 1/\beta)$).

Note that $t \geq \alpha \Rightarrow v \geq \alpha$.

$S_2 \leq (\frac{1}{2})^{3n} \sum_{u=\alpha}^{n-\alpha} \binom{n}{u} \sum_{s=1}^{\alpha} n^{\alpha} \sum_{t=\alpha}^v \binom{v}{t} (3(\frac{1}{2})^{\alpha})^{m-3} = o((\frac{1}{2})^{2n})$ for large enough α .

$S_3 \leq (\frac{1}{2})^{3n} \sum_{u=1}^{\alpha} n^{2\alpha} \sum_{s=1}^{\alpha} \sum_{t=\alpha}^v \binom{v}{t} ((\frac{1}{2})^{\alpha} + (\frac{1}{2}) + (\frac{1}{2})^{\alpha})^{m-3} = o((\frac{1}{2})^{2n})$ for large enough α .

$S_4 \leq (\frac{1}{2})^{3n} \alpha^3 n^{3\alpha} = o((\frac{1}{2})^{2n})$.

Consequently $P(c \not\rightarrow b \not\rightarrow a) - 3(\frac{1}{2})^{2n} = o((\frac{1}{2})^{2n})$. \square

(v) $d \not\rightarrow b \not\rightarrow a$

Lemma 3.5. $P(d \not\rightarrow b \not\rightarrow a) \sim (\frac{1}{2})^{m+n-2} + (\frac{1}{2})^{2n}$.

Proof. Let $X' = X - \{a, b\}$, $Y' = Y - \{d\}$.

As before, we have the simple lower bound: $P(d \not\rightarrow b \not\rightarrow a) \geq P(\{O_d = O_b = \emptyset\} \cup \{O_d = I_a = \emptyset\} \cup \{I_b = I_a = \emptyset\}) \geq (\frac{1}{2})^{m+n-2} + (\frac{1}{2})^{2n} - (\frac{1}{2})^{m+2n-3}$.

We bound the probability from above by the probability of there being no path from d to b or from b to a of length at most 3 or 4 respectively. The edges $\{a, d\}$ and $\{b, d\}$ are both directed towards d . Condition on $O_d = T$, $I_b = U$, $I_a = S$, $O_b = V$. No edge is directed from T to U , and $S \subseteq U$. These conditions imply that no $x \in S$ is a VU -witness. In addition we forbid any $x \in X' - T$ to be a VS -witness. The events ' x is a VS -witness' are independent for $x \in X' - T$ and independent of the other necessary events just stated. We obtain

$$P(d \not\rightarrow b \not\rightarrow a)$$

$$= \sum_{S, T, U, V} P(d \not\rightarrow b \not\rightarrow a | O_d = T, I_b = U, O_b = V, I_a = S) P(O_d = S, I_b = T, I_a = U, I_a = S)$$

$$\begin{aligned}
&\leq \left(\frac{1}{2}\right)^{2n+m-2} \sum_{t=0}^{m-2} \binom{m-2}{t} \sum_{u=0}^{n-1} \binom{n-1}{u} \sum_{s=0}^u \binom{u}{s} \left(\frac{1}{2}\right)^{st} \left(\left(\frac{1}{2}\right)^s + \left(\frac{1}{2}\right)^v - \left(\frac{1}{2}\right)^{s+v} \right)^{m-2-t} \\
&= \left(\frac{1}{2}\right)^{2n+m-2} \sum_{u=0}^{n-1} \sum_{s=0}^u \binom{n-1}{u} \binom{u}{s} \left(\left(\frac{1}{2}\right)^u + \left(\frac{1}{2}\right)^v + \left(\frac{1}{2}\right)^s - \left(\frac{1}{2}\right)^{s+v} \right)^{m-2}.
\end{aligned}$$

For $s = 0$ we obtain the following sum: $\left(\frac{1}{2}\right)^{2n+m-2} \sum_{u=0}^{n-1} \binom{n-1}{u} \left(\left(\frac{1}{2}\right)^u + 1 \right)^{m-2} = \left(\frac{1}{2}\right)^{2n+m-2} \left(2^{m-2} + (n-1)^\alpha \alpha \left(\frac{3}{2}\right)^{m-2} + \left(1 + \left(\frac{1}{2}\right)^\alpha\right)^{m-2} \right) = \left(\frac{1}{2}\right)^{2n} + (n-1)^\alpha \alpha \left(\frac{1}{2}\right)^{2n} \left(\frac{3}{2}\right)^{m-2} + \left(\frac{1}{2}\right)^{2n+m-2} \left(1 + \left(\frac{1}{2}\right)^\alpha\right)^{m-2} = \left(\frac{1}{2}\right)^{2n} + o\left(\left(\frac{1}{2}\right)^{m+n-2} + \left(\frac{1}{2}\right)^{2n}\right).$

For $v = 0 = n - 1 - u$: $\left(\frac{1}{2}\right)^{2n+m-2} \sum_{s=0}^{n-1} \binom{n-1}{s} \left(1 + \left(\frac{1}{2}\right)^{n-1}\right)^{m-2} = \left(\frac{1}{2}\right)^{m+n-2} \left(1 + \left(\frac{1}{2}\right)^{n-1}\right)^{m-2} = \Theta\left(\left(\frac{1}{2}\right)^{m+n-1}\right)$, as $\lim_{n \rightarrow \infty} \left(1 + \left(\frac{1}{2}\right)^{n-1}\right)^{m-2} = 1$.

Split the remaining sum into the following cases: $S_1 : 1 \leq u \leq \alpha$, $S_2 : 1 \leq s \leq \alpha \leq u, v$, $S_3 : \alpha \leq s, u, v$, $S_4 : 1 \leq s, v \leq \alpha$, $S_5 : 1 \leq v \leq \alpha \leq s$.

Clearly, $S_1 = o\left(\left(\frac{1}{2}\right)^{m+2n-2} + \left(\frac{1}{2}\right)^{2n}\right)$.

$S_2 \leq n^\alpha \left(\frac{1}{2}\right)^{m+2n-2} \sum_{u=\alpha}^{n-1} \binom{n-1}{u} \left(2 \left(\frac{1}{2}\right)^\alpha + \left(\frac{1}{2}\right)^s\right)^{m-2} = o\left(\left(\frac{1}{2}\right)^{m+2n-2} + \left(\frac{1}{2}\right)^{2n}\right)$, for $\alpha > 21$, which is easily seen by considering $\beta \leq 1$ and $\beta > 1$ separately.

$S_3 \leq \left(\frac{1}{2}\right)^{m+2n-2} \sum_{u=\alpha}^{n-1} \binom{n-1}{u} \sum_{s=\alpha}^u \binom{u}{s} \left(3 \cdot \left(\frac{1}{2}\right)^\alpha\right)^{m-2} \leq \left(\frac{1}{2}\right)^{m+2n-2} \left(\frac{3}{2^\alpha}\right)^{m-2} \cdot 3^{n-1} = o\left(\left(\frac{1}{2}\right)^{m+n-2} + \left(\frac{1}{2}\right)^{2n}\right)$.

$S_4 = o\left(\left(\frac{1}{2}\right)^{m+2n-2} + \left(\frac{1}{2}\right)^{2n}\right)$, since S_4 is the sum of a constant (α^2) number of $o\left(\left(\frac{1}{2}\right)^{m+2n-2} + \left(\frac{1}{2}\right)^{2n}\right)$ terms.

$S_5 \leq \alpha n^\alpha \left(\frac{1}{2}\right)^{m+2n-2} \sum_{s=\alpha}^{n-1} \binom{n-1}{s} \left(\left(\frac{1}{2}\right)^\alpha + \left(\frac{1}{2}\right)^v + \left(\frac{1}{2}\right)^\alpha\right)^{m-2} = \alpha n^\alpha \left(\frac{1}{2}\right)^{m+n} \left(\left(\frac{1}{2}\right) + \left(\frac{1}{2}\right)^{\alpha-1}\right)^{m-2} = o\left(\left(\frac{1}{2}\right)^{m+2n-2} + \left(\frac{1}{2}\right)^{2n}\right)$ for $\alpha > 2$.

Finally,

$$P(d \nrightarrow b \nrightarrow a) \sim \left(\frac{1}{2}\right)^{m+n-2} + \left(\frac{1}{2}\right)^{2n}.$$

□

We now show how to use the lemmas above to prove Theorem 1. For example, suppose $\beta < 1$, $x, y \in X$, and $z \in Y$. Then

$$\begin{aligned}
RC_{m,n} &= 1 - \frac{P(x \nrightarrow y)P(y \nrightarrow z)}{P(x \nrightarrow y \nrightarrow z)} = \\
1 - \frac{P(a \nrightarrow b)P(b \nrightarrow d)}{P(a \nrightarrow b \nrightarrow d)} &= 1 - \frac{P(b \nrightarrow a)P(d \nrightarrow a)}{P(d \nrightarrow b \nrightarrow a)} = \text{(by lemmas 3.5, 3.1 and 3.2)} \\
&= 1 - \frac{2 \left(\frac{1}{2}\right)^n \left(\left(\frac{1}{2}\right)^m + \left(\frac{1}{2}\right)^n\right)}{4 \left(\frac{1}{2}\right)^{m+n} + \left(\frac{1}{2}\right)^{2n}}.
\end{aligned}$$

Now, since $\beta < 1$ we get

$$\lim_{n \rightarrow \infty} RC_{m,n} = \lim_{n \rightarrow \infty} 1 - \frac{2(1 - \left(\frac{1}{2}\right)^{n-m})}{4 + \left(\frac{1}{2}\right)^{n-m}} = \frac{1}{2}.$$

The other entries in the table in the statement of the theorem can be found similarly.

4 Further questions

I would like to mention two questions:

Question 1:

How do the results above change if an edge $e = \{x, y\}$, where $x \in X$ and $y \in Y$, rather than being oriented either way with equal probability, is directed from X to Y with some fixed probability p ?

Question 2:

From Theorem 1, it seems plausible that $\{x \nrightarrow y\}$ and $\{y \nrightarrow z\}$ should be negatively correlated in any complete bipartite graph (at least in any large enough graph) when x, y, z belong to the same color class. This seems not to be the case, however; computer calculations show that if n is much larger than m (on the scale $n = 2^m$), then the events mentioned seem to be positively correlated even for fairly large n . Is this true in general? Can the cases with positive correlation be completely identified in terms of m and n ?

The calculations mentioned above made use of the following recursions:

$$\begin{aligned} f_X(m, n, k) &= \sum_{l=0}^n \binom{n}{l} \frac{(2^k - 1)^l}{2^{nk}} f_Y(m - k, n, l), \\ f_Y(m, n, l) &= \sum_{k=0}^{m-1} \binom{m-1}{l} \frac{(2^k - 1)^l}{2^{mk}} f_X(m, n - l, k), \\ g_X(m, n, k) &= \sum_{l=0}^n \binom{n}{l} \frac{(2^k - 1)^l}{2^{nk}} g_Y(m - k, n, l), \\ g_Y(m, n, l) &= \sum_{k=0}^{m-2} \binom{m-2}{k} \frac{(2^l - 1)^k}{2^{lm}} g_X(m, n - l, k), \\ h_X(m, n, k) &= \sum_{l=0}^{n-1} \binom{n-1}{l} \frac{(2^k - 1)^l}{2^{nk}} h_Y(m - k, n, l), \end{aligned}$$

and

$$h_Y(m, n, l) = \sum_{k=0}^{m-1} \binom{m-1}{k} \frac{(2^l - 1)^k}{2^{mk}} h_X(m, n - l, k).$$

where f_X, f_Y, g_X, g_Y, h_X and h_Y are defined as follows.

Let $P_{m,n}$ denote the probability measure associated with a uniformly chosen orientation of $K_{m,n}$, where the class X has size m and the class Y has size n . For a subset K of vertices of $K_{m,n}$ and a vertex x in $K_{m,n}$, let $\{K \nrightarrow x\} = \bigcap_{k \in K} \{k \nrightarrow x\}$. Let $a, b, c \in X, d \in Y$ and K be any subset of X , not including a or b , of size k and L be any subset of Y , not including d , of size l .

Then $f_X(m, n, k) = P_{m,n}(K \nrightarrow a)$, $f_Y(m, n, l) = P_{m,n}(L \nrightarrow a)$, $g_X(m, n, k) = P_{m,n}(K \nrightarrow b \text{ and } b \nrightarrow a)$, $g_Y(m, n, l) = P_{m,n}(L \nrightarrow b$

and $b \nrightarrow a$, $h_X(m, n, k) = P_{m,n}(K \nrightarrow d \text{ and } d \nrightarrow a)$, and $h_Y(m, n, l) = P_{m,n}(L \nrightarrow b \text{ and } b \nrightarrow a)$.

In addition, we have the following base cases for the formulas above: $g_X(m, n, 0) = g_Y(m, n, 0) = f_X(m, n, 1)$, $h_X(m, n, 0) = h_Y(m, n, 0) = f_Y(m, n, 1)$, and $f_X(m, n, 0) = f_Y(m, n, 0) = 1$.

These functions are related to the quantities estimated in the lemmas above by $f_X(m, n, 1) = P_{m,n}(a \nrightarrow b)$, $f_Y(m, n, 1) = P_{m,n}(a \nrightarrow d)$, $g_X(m, n, 1) = P_{m,n}(c \nrightarrow b \nrightarrow a)$, $g_Y(m, n, 1) = P(d \nrightarrow b \nrightarrow a)$, and $h_X(m, n, 1) = P(c \nrightarrow d \nrightarrow a)$.

Acknowledgement

I thank Svante Linusson for suggesting the problem and for helpful suggestions on earlier versions of the manuscript.

References

- [1] Sven Erick Alm and Svante Linusson, A counter-intuitive correlation in a random tournament (2009). To appear in *Comb. Prob. and Computing*.
- [2] Sven Erick Alm and Svante Linusson, Correlations for paths in random orientations of $G(n, p)$ (2009), arXiv:0906.0720.
- [3] Theodore E. Harris, A lower bound for the critical probability in a certain percolation process, *Proc. Cambridge Philos. Soc.* **58** 13-20 (1960).
- [4] Colin McDiarmid, General Percolation and random graphs *Adv. Appl. Prob.* **13** 40-60 (1981).

Department of Mathematics, KTH - Royal Institute of Technology, SE-100 44 Stockholm, Sweden.

E-Mail: `eaas@math.kth.se`.